Voxel-Based Irregularity Age Map (IAM) for Brain's White Matter Hyperintensities in MRI

Voxel-Based Irregularity Age Map (IAM) for Brain's White Matter Hyperintensities in MRI

Muhammad Febrian Rachmadi*†, Maria del C. Valdés-Hernández†, Taku Komura*,
and The Alzheimer's Disease Neuroimaging Initiative[1]
*School of Informatics, The University of Edinburgh, Edinburgh, UK
†Centre for Clinical Brain Sciences, The University of Edinburgh, Edinburgh, UK
email: febrian.rachmadi@ed.ac.uk

***Abstract*—In this paper, we propose a novel way to produce voxel-based irregularity age map (IAM) for brain magnetic resonance image (MRI) to identify white matter hyperintensities (WMH) on scans with mild vascular pathology. Age map is a term used in computer graphic field that reveals age/progression of defected areas in images' texture. In this work, age map is used to reveal the age of irregularity of brain tissue (*i.e.*, hyperintensities). Age map of WMH is useful because it shows not only the probability of voxels to be WMH but also the scale in the progression of voxels to become WMH. Our approach is fully automatic and unsupervised with little to none human interaction. We evaluated our approach using brain MRI data obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database and visually compared the results with those obtained from the public toolbox Lesion Segmentation Toolbox (LST). We also evaluated our proposed approach on images from 10 different subjects using Dice similarity coefficient (DSC).**

I. Introduction

White matter hyperintensities (WMH) are regions in brain magnetic resonance image (MRI) that appear brighter than the other regions in the T2-Fluid Attenuation Inversion Recovery (FLAIR) MRI sequence. WMH are features found in many brain pathologies such as stroke, multiple sclerosis (MS) and dementia/Alzheimer's Disease (AD). Specifically, it is believed that WMH are associated with the progression of dementia [1][2]. Because of that, producing age map of WMH is beneficial and important to track the progression of brain pathologies.

In past years, there have been many attempts to develop WMH segmentation tools where most of works use machine learning algorithms. Support vector machine (SVM)[3], random forest (RF)[4] and convolution encoder network (CEN)[5] are some supervised machine learning methods that have been studied for automatically segmenting WMH. However, supervised methods are limited to the availability of labelled data, which is very time consuming to produce and highly dependent on physician or expert's delineation. To avoid the aforementioned dependencies, unsupervised methods have been developed, tested and available as public toolboxes such as Lesion Segmentation Tool (LST) [6] and Lesion-TOADS [7]. Unfortunately, performances of these toolboxes, which employ unsupervised methods, are still limited compared to supervised methods [4][5].

We propose a novel way to extract irregularity age map that shows age of voxels as hyperintensities instead of probability of voxels being hyperintensities. The irregularity age map proposed in this paper is a map where each brain voxel has value from 0 to 1 to indicate a scale in the progression of irregularities, *i.e.*, hyperintensities, from their minimal up to their maximal intensity degree. In this age map, brighter intensities are indicated with higher value (*i.e.*, close to 1) than the others (*i.e.*, close 0). This differs from probability maps where each voxel has a probability value to indicate whether the voxel is WMH or not. This approach is derived from a previously published computer graphic work [8], where they extract the age map to indicate weathered area of some texture images.

Age map calculation and generation have some benefits compared to other approaches. Firstly, the age map calculation and generation do not need hyperintensities to be labelled, which are very tedious and expensive to produce. Secondly, the age map of MRI shows all irregularities inside the brain, not only one from a specific brain pathology. In this work, we use hyperintensities as irregularities of MR images. However, it is possible to generate age maps for different irregularities such as lacunar lesions or traumatic injuries.

This paper is organised as follow. Section 1 discusses some backgrounds of this study. Section 2 discusses related work and overview of our proposed approach. Section 3 explains step-by-step on how irregularity age maps are calculated and generated. Section 4 and Section 5 show some results for visual observation and its discussions. Section 6 draws conclusions of this study and some future works.

II. Related Work and Proposed Approach

This work is highly influenced from previous work in computer graphic [8] where authors successfully generate age map from texture images. The main idea of age map is to quantify irregularities on weathered texture images into age values in the range of 0 to 1, where weathered areas (*i.e.*, consumed by time and weather) are regarded as irregularities

[1]Data used in preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

of the texture. The age map is produced by calculating distance values between *source patches* and *target patches*. Source patches are made by dividing the texture image into grid of $N \times N$ non-overlapping patches. Whereas, target patches are all possible patches of the same size of the source patch extracted from the same texture image. A distance value is calculated by computing the difference of luminance and gradient's direction between two patches: a source patch and a target patch. Lower distance value means higher similarity between two patches, while higher distance value means lower similarity. All quantified source patches make a complete age map of texture image where hot/red areas (*i.e.*, age value near to 1) indicate severe irregularities.

The same approach of producing age map can be applied for brain MRI because hyperintensities are also tissue irregularities in MRI. However, some prerequisites and modifications in the calculation of the distance value are needed. In our proposed approach, not only the original MR image but also intra-cranial volume (ICV) and cerebrospinal fluid (CSF) masks are needed. The masks are used to mask out non-brain tissue voxels for the age map calculation. After that, hierarchical patch creation with different patches size and patch generation of source patches and target patches are performed. Finally, age values are calculated using distance function and nearest neighbour. Age map generation, which does global normalisation and penalty to the age value calculation results, is also used to produce the final age map. All these steps will be thoroughly discussed in the next section.

III. Age Map Calculation and Generation

In this section, all steps of the proposed approach are discussed. The steps are divided in four and described in subsections. These are: data preparation (*i.e.*, brain MRI and its masks), patch generation (*i.e.*, hierarchical patch generation for source and target patches), age value calculation (*i.e.*, distance value, nearest neighbour approach and hierarchical age value combination) and final age map generation (*i.e.*, global normalisation and penalty).

A. Data Preparation: Masks of Brain MRI

The original brain MRI includes non-brain tissues that are not needed in the age map calculation such as skull, cerebrospinal fluid, veins and meninges. The inclusion of these regions can make computation slower as more unrelated voxels would be included in the computation, and can bias the age value computation because brain voxels would be compared to non-brain voxels. Masks used in this study (*i.e.*, ICV mask and CSF mask) are produced using an in-house pipeline developed by the University of Edinburgh. The ICV mask is a brain mask that excludes skull and background from brain MRI. Whereas, CSF is a mask which extracts brain tissue from cerebrospinal fluid, veins and meninges. Example of original brain MRI overlaid by ICV and CSF masks can be seen in Fig. 1.

B. Patch Generation: Source Patch and Target Patch

Like in previous work [8], the original MR image is divided into non-overlapping grid-patches to create *source patches*.

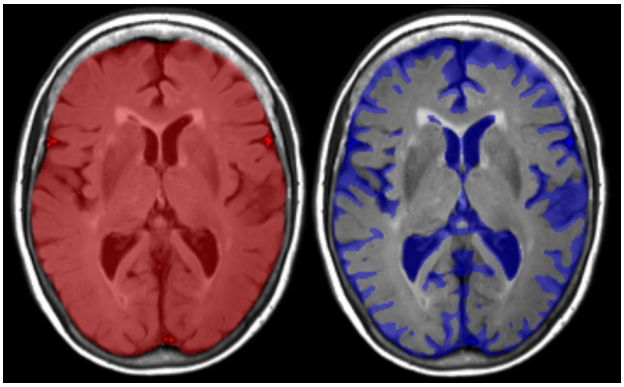


Fig. 1: Original brain MRI overlaid by intra-cranial volume (ICV) mask (red) and cerebrospinal fluid (CSF) mask (blue).

However, in this study, we create source patches using four different sizes (*i.e.*, 1×1, 2×2, 4×4 and 8×8). Note that the division is done slice by slice, not in the whole MRI (3D) scan. For each source patch, we then generate random same-sized patches sampled from all possible patches of the same slice for *target patches*. Fig. 2 visualises the patch generation for source patches and target patches. Unlike in previous work [8], we do not use all possible patches available for target patches to fasten computation in the next stage. Preliminary pilot tests done in our sample showed no indication that this affects quality of the age map. Furthermore, we skip patch generation and age map calculation in slices without or too few brain voxels to fasten the computation even more.

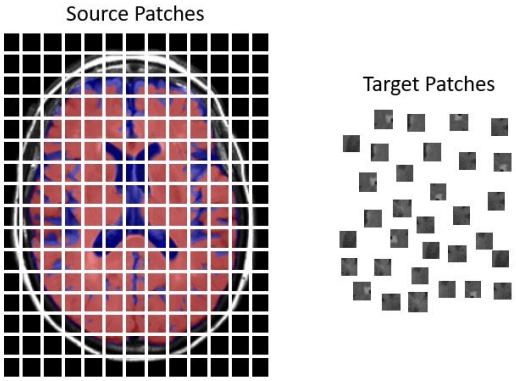


Fig. 2: Patch generation visualised. Source patch is included in computation if it is inside of the brain and not cerebrospinal fluids (*i.e.*, red area). Target patches are generated with the same size as source patch for all possible overlapping patches inside red area from the same slice.

We use four different sizes of patches, which are 1×1, 2×2, 4×4 and 8×8, to make sure we capture different details of brain tissue. Smaller patches capture more detailed information from brain MRI while bigger patches capture bigger contextual information. Each of them are computed independently, so there will be four different age maps for each MRI slice. They will be combined together in the last step of age value

calculation to produce the final MRI age map. The process is repeated sequentially across all MRI slices.

To speed-up computation, we do some exclusions in patch generation. In [8], each source patch is compared to all possible overlapping patches from the same image. However, in this study, we do not use all possible overlapping patches for target patches, but they are randomly sampled from all possible overlapping brain patches. Non-brain patches are excluded from random sampling using masks. Moreover, we also exclude neighbouring target patches of source patches because neighbouring patches are too similar (also done in [8]). Lastly, we only generate target patches if the centre of the source patch is within the ICV mask and outside the CSF mask (i.e. in brain tissue). In general, there are 300-1300 target patches sampled for each source patch.

C. Age Value Calculation: Distance Function

The core computation of the age map generation is age value calculation using a distance function. Age value calculation is done by comparing how distant/different source patch and target patch are, and a distance function is needed for computation. Let $\mathbf{s}$ be a source patch and $\mathbf{t}$ a target patch, then the age/distance value of the two patches d is:

$$d = \alpha \cdot |\text{max}(\mathbf{s} - \mathbf{t})| + (1 - \alpha) \cdot |\text{mean}(\mathbf{s} - \mathbf{t})| \,. \quad (1)$$

According to Equation 1, the age value is computed by subtracting the target patches intensities out from the source patches intensities, calculating its absolute maximum and mean values, and then adding them together with a parameter of α (*e.g.*, $\alpha = 0.5$ in this study). The rationale behind calculating this distance function is that we want to know the maximum difference between source and target patches relative to the mean difference of the two patches. Note that both patches can be arrays/matrices in which we can find theirs maximum and mean values.

To calculate the age value of one particular source patch, we need to find its nearest neighbour based on a list of age values from a list of target patches generated from the last step. This can be done by sorting all age values calculated from all target patches in ascending order, and finding the mean age value from the 100 biggest age values in the sorted list. The rationale behind this is that if a source patch is similar to the majority of target patches, then the mean value of the 100 biggest age values of the sorted list should be still lower than other source patches that are very different than the majority of their target patches. After all age values for all source patches in one slice have been calculated, age value normalisation to $[0, 1]$ can be done to create the age map of an MRI slice.

From a previous step, we will have four different age maps of MRI slice from four different sizes of patches. We can combine all of them by using the following formulation:

$$cMRI = \alpha \cdot map_1 + \beta \cdot map_2 + \gamma \cdot map_4 + \delta \cdot mri_8 \quad (2)$$

where $\alpha + \beta + \gamma + \delta = 1$ and map_1, map_2, map_4 and map_8 stand for age map from different size of patch, which are 1×1, 2×2, 4×4 and 8×8 respectively. Equation 2 is basically a weighted blend of four different age maps. Also, please note that the size of the age map of an MRI slice from different patch sizes differ depending on the size of the non-overlapping grid patches (*e.g.*, if the patches are of size 1×1 we will have a 256×256 grid in the MRI slice, and, thus, we will have a 256×256 age map. If we use patches of size 2×2, we will have a 128×128 age map. If we use patches of size 4×4, we will have a 64×64 age map, and so on.). To make all age maps of the same size, the age maps are up-sampled to fit the original size of the MRI slices (*i.e.*, 256×256 in this study). We also apply Gaussian blur filter to the resized age maps to make them less pixelated. Example results from these steps can be seen in Fig. 3.

D. Final Age Map: Penalty and Global Normalisation

The age map generated enhanced voxels that are hypo- and hyper-intense with respect to the normal tissue by means of the absolute intensity difference, as per Equation 1. This means that, at this point, we cannot differentiate between hyper-intensities (*i.e.*, bright voxels) and hypo-intensities (*i.e.*, dark voxels). Most hypo-intensities are excluded using ICV and CSF masks, but there is still possibility that they appear and make some disturbing artefacts. Moreover, the probability of artefacts to appear is much higher if we do not have high quality masks. To subdue these artefacts and refine age map result, we use the following penalty formulation:

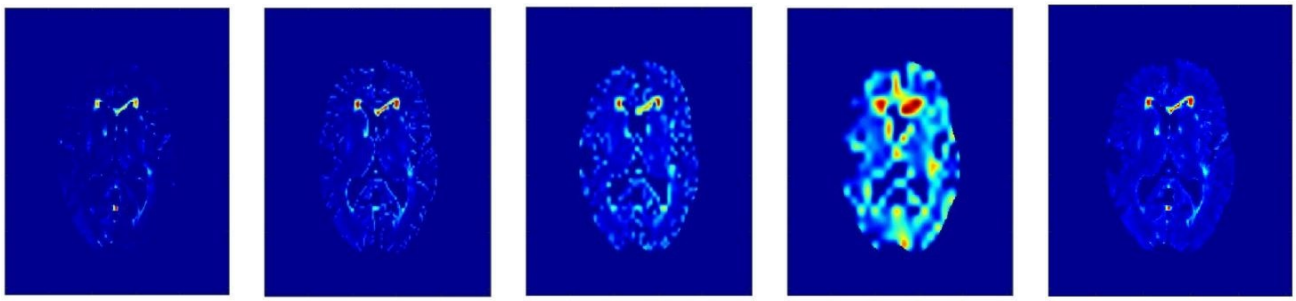
$$p_o = p_i \times i \quad (3)$$

where p_i is voxel from resulted age map, i is voxel's intensity of the original MRI and p_o is output voxel after penalty. In general term, we do penalty by multiplying all age map's voxels with theirs correspondent original intensity of MRI.

The last step in the generation of the final age map is global normalisation (GN) where we do normalisation to $[0, 1]$ values for one MRI data (*i.e.*, all slices together). Until now, we do normalisation per slice which is limited to local voxels. GN is useful to subdue biases that usually come differently from each slice, especially from the cerebellum. Furthermore, biases which can appear when there is no hyperintensities in a slice are also subdued by the GN. Example of GN and penalty results are depicted in Fig. 4.

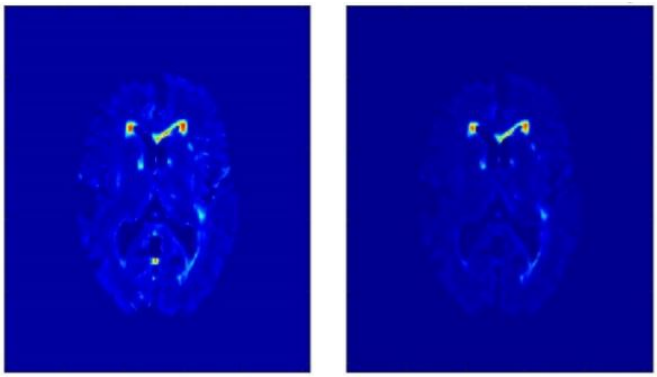
IV. Experiment and Visual Results

We used T2-Fluid Attenuation Inversion Recovery (FLAIR) MRI data from Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. Since not all data in this dataset have the corresponding WMH labels available, two different evaluations are done. Firstly, we compare the results from the proposed approach with those from a public toolbox, which is Lesion Growth Algorithm from Lesion Segmentation Tool (LST-LGA) with parameter $\kappa = 0.3$ (*i.e.*, default value of parameter). We choose LST-LGA because it is an unsupervised approach, similar to our proposed approach. Some examples for visual observation can be seen in Fig. 6. Secondly, we calculate the similarity index Dice coefficient (DSC) [9] from 20 different subjects for which the corresponding WMH labels



(a) (b) (c) (d) (e)

Fig. 3: Example results of age maps of MRI slice. Fig. 5a, 5b, 3c and 3d are generated from different sizes of patch, which are 1×1, 2×2, 4×4 and 8×8 respectively, while Fig. 3e is the combination of all four age maps mentioned before. The original MRI and its masks are depicted in Fig. 1. Please refer to the digital version of this paper to see in more detail.



(a) (b)

Fig. 4: Example of (a) global normalisation (GN) and (b) Combined use of GN and penalty to age map. The original MRI and its masks can be seen in Fig. 1, while its age map before GN and penalty is depicted in Fig. 3e.

and all masks needed for computation are available. DSC scores show how similar/dissimilar segmentation results are with respect to the reference labels (*i.e.*, groundtruth), and are calculated using the following equation:

$$DSC = \frac{2 \times TP}{FP + 2 \times TP + FN} \tag{4}$$

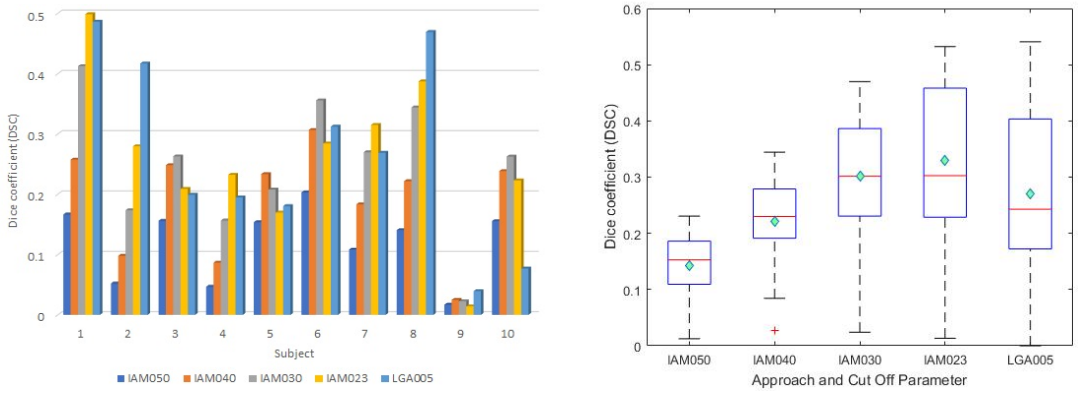
where TP is true positive, FP is false positive and FN is false negative. Higher DSC score means better performance. To produce WMH segmentations from the final age map, arbitrary threshold values can be used. In this study, we used four different thresholds which are 0.50, 0.40, 0.30 and 0.23. We compare the DSC scores from the results of the proposed method against those from the LST-LGA [5] with parameter $\kappa = 0.05$. The results are depicted in Fig. 5, where a bar chart shows performance in 10 random subjects and box-plots show distribution of performances for each approach.

V. Discussion

DSC evaluations for WMH segmentation in Fig. 5 show that IAM perform generally better than LST-LGA. The DSC scores of IAM with threshold = 0.23 are higher than LST-LGA as shown in box-plot (average performances = 0.3301 and 0.2698 respectively). Furthermore, threshold = 0.23 yields better performance in IAM than 0.50, 0.40 and 0.30 threshold values. On the other hand, visual observation can be seen in Fig. 6 where hot area (red) means age value is higher while cold area (blue) means age value is lower. It can be observed that our proposed approach IAM works well in most cases, especially when there is enough number of WMH in a slice. However, it still struggles when there is limited number of WMH in a slice as shown in the fourth row of Fig. 6, which is the case of a healthy subject. In this case, our approach gives high irregularity age values to regions of the brain cortex. Furthermore, another shortcoming that can be seen in Fig. 6 is that some subtle hyperintensities do not have high age values compared to other voxels, which are depicted in the first and third row. We believe that this is caused by the penalty term as subtle hyperintensities do not have as high intensity values as strong hyperintensities. However, our proposed approach is generally better than LST-LGA visually.

VI. Conclusion

In this study, we successfully implemented a voxel-based irregularity age map generation from T2-FLAIR brain MRI. Compared to LST-LGA, also an unsupervised WMH segmentation method, our proposed approach is visually better for both strong hyperintensities and subtle hyperintensities. Some shortcomings still can be seen in the cortex. Our proposed approach depends on the existent of hyperintensities: it struggles if there are no hyperintensities at all in a slice. Some weak hyperintensities are wrongly subdued due to the penalty term imposed. However, we believe that better normalisation and penalty terms can eliminate these two problems in future studies. Another minor shortcoming is that our proposed approach highly depends on pre-processing and post-processing. High quality masks, normalisation term and penalty terms are needed to exclude non-relevant regions before computation and subdue artefacts.



(a) (b)

Fig. 5: Dice coefficient (DSC) for 10 random subjects from different approaches (a) and distribution of DSC for each approach (b) where higher score of DSC means better performance. IAM stands for Irregularity Age Map (*i.e.*, proposed approach), and the following numbers stands for threshold value t for WMH segmentation (*i.e.*, 050 means $t = 0.50$). LGA005 stands for Lesions Segmentation Tool's Lesion Growth Algorithm with parameter $\kappa = 0.05$. Red lines in the box-plot stand for median of each approach (*i.e.*, 0.1529, 0.2300, 0.3015, 0.3027 and 0.2429 respectively) whereas green diamonds stand for mean of each approach (*i.e.*, 0.1427, 0.2122, 0.3011, 0.3301 and 0.2698 respectively)

ACKNOWLEDGMENT

The first author would like to thank Indonesia Endowment Fund for Education (LPDP) of Ministry of Finance, Republic of Indonesia, for funding his study at School of Informatics, the University of Edinburgh. Funds from Row Fogo Charitable Trust (MCVH) are also gratefully acknowledged. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimers Association; Alzheimers Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research and Development, LLC.; Johnson and Johnson Pharmaceutical Research and Development LLC.; Lumosity; Lundbeck; Merck and Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimers Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

REFERENCES

[1] J. M. Wardlaw, E. E. Smith, G. J. Biessels, C. Cordonnier, F. Fazekas, R. Frayne, R. I. Lindley, J. T O'Brien, F. Barkhof, O. R. Benavente *et al.*, "Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration," *The Lancet Neurology*, vol. 12, no. 8, pp. 822–838, 2013.

[2] A. C. Birdsill, R. L. Koscik, E. M. Jonaitis, S. C. Johnson, O. C. Okonkwo, B. P. Hermann, A. LaRue, M. A. Sager, and B. B. Bendlin, "Regional white matter hyperintensities: aging, alzheimer's disease risk, and cognitive function," *Neurobiology of aging*, vol. 35, no. 4, pp. 769–776, 2014.

[3] S. Klöppel, A. Abdulkadir, S. Hadjidemetriou, S. Issleib, L. Frings, T. N. Thanh, I. Mader, S. J. Teipel, M. Hüll, and O. Ronneberger, "A comparison of different automated methods for the detection of white matter lesions in mri data," *NeuroImage*, vol. 57, no. 2, pp. 416–422, 2011.

[4] V. Ithapu, V. Singh, C. Lindner, B. P. Austin, C. Hinrichs, C. M. Carlsson, B. B. Bendlin, and S. C. Johnson, "Extracting and summarizing white matter hyperintensities using supervised segmentation methods in alzheimer's disease risk and aging studies," *Human brain mapping*, vol. 35, no. 8, pp. 4219–4235, 2014.

[5] M. Rachmadi, T. Komura, M. Valdes Hernandez, and M. Agan, *Evaluation of Four Supervised Learning Schemes in White Matter Hyperintensities Segmentation in Absence or Mild Presence of Vascular Pathology*, 4 2017.

[6] P. Schmidt, C. Gaser, M. Arsic, D. Buck, A. Förschler, A. Berthele, M. Hoshi, R. Ilg, V. J. Schmid, C. Zimmer *et al.*, "An automated tool for detection of flair-hyperintense white-matter lesions in multiple sclerosis," *Neuroimage*, vol. 59, no. 4, pp. 3774–3783, 2012.

[7] N. Shiee, P.-L. Bazin, A. Ozturk, D. S. Reich, P. A. Calabresi, and D. L. Pham, "A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions," *NeuroImage*, vol. 49, no. 2, pp. 1524–1535, 2010.

[8] R. Bellini, Y. Kleiman, and D. Cohen-Or, "Time-varying weathering in texture space," *ACM Transactions on Graphics (TOG)*, vol. 35, no. 4, p. 141, 2016.

[9] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

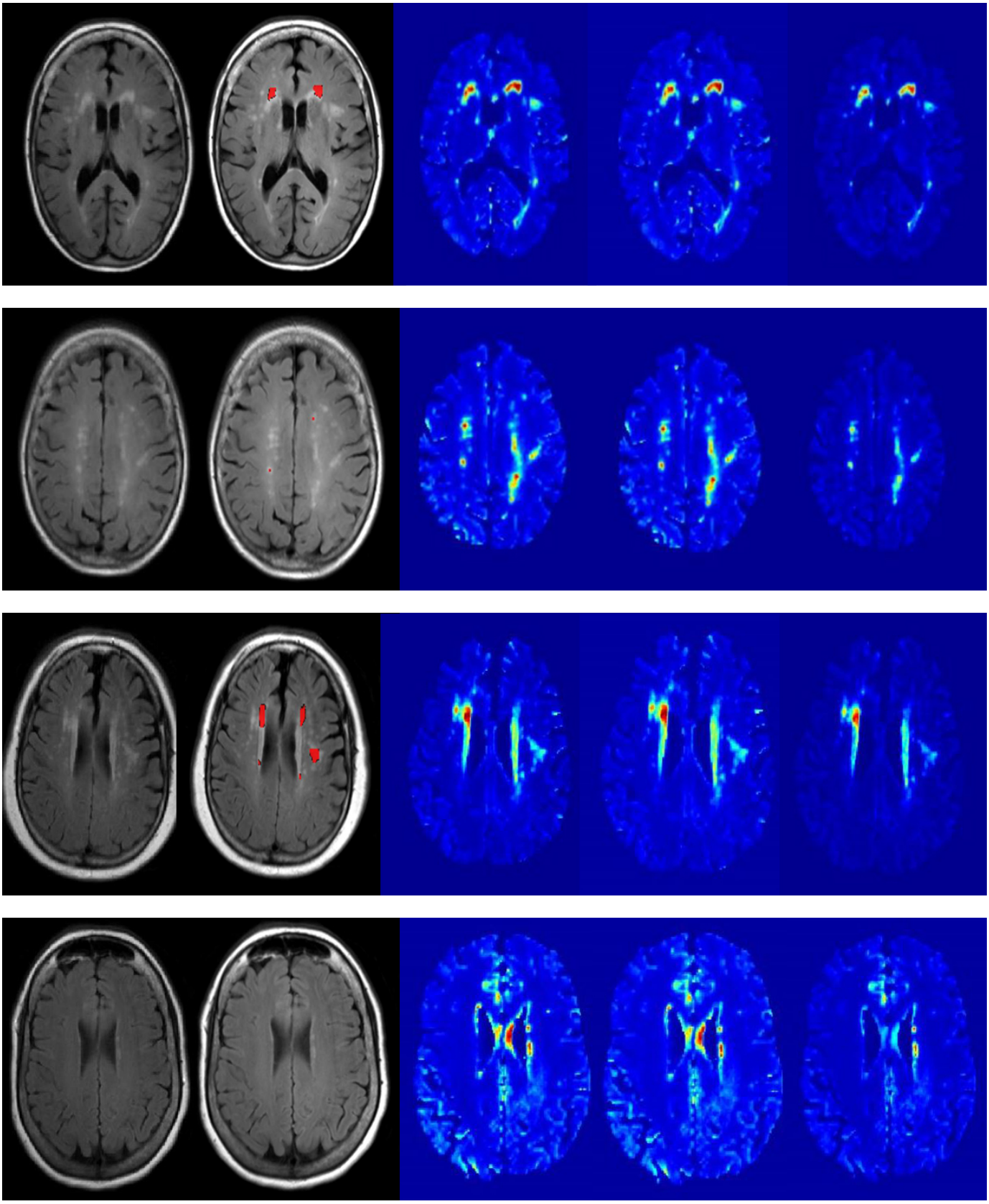


Fig. 6: Original MRI data and results from different toolbox and steps of our proposed approach. From the left respectively are original MRI data, Lesion Growth Algorithm of Lesion Segmentation Toolbox (LST-LGA), combined age map from four different sizes of patch, combined age map with global normalisation, and combined age map with global normalisation and penalty. Red regions on LST-LGA results are WMH automatically segmented by LST-LGA with parameter $\kappa = 0.3$ (*i.e.*, default value of parameter).